

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2002, 09:41:26 ; Search time 517.9 Seconds
(without alignments)
5944.058 Million cell updates/sec

Title: US-09-719-017a-2

Perfect score: 1793

Sequence: 1 gattccccgttgacaatla.....caattactcatgcgcgcg 1793

Scoring table: IDENTITY_NUC

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1793	100.0	1793	21	AAA47190
2	1793	100.0	1793	21	AAZ45325
3	1126.4	62.8	1727	15	AAO86279
4	954.2	53.2	1071	21	AAA12173
5	404	22.5	5096	16	AAO83824
6	404	22.5	5110	16	AAO83833
7	404	22.5	5110	16	AAO83834
8	404	22.5	5110	16	AAO83835
9	404	22.5	5110	16	AAO83836

10	404	22.5	5110	16	AAO83825	Plasmid pcm-x# S
11	404	22.5	5110	16	AAO83826	Plasmid pcAT10-1.
12	404	22.5	5110	16	AAO83827	Plasmid pcm-T297.
13	404	22.5	5110	16	AAO83828	Plasmid pcm-p1160.
14	404	22.5	5110	16	AAO83829	Plasmid pcm-T7270.
15	404	22.5	5110	16	AAO83830	Plasmid pcm-P1087.
16	404	22.5	5110	16	AAO83831	Plasmid pcm-P1198.
17	404	22.5	5110	16	AAO83832	Plasmid pcm-T286.
18	403.6	22.5	502	16	AAO92246	rrnB ribosomal RNA
19	401.6	22.4	7615	20	AAO83833	Telomerase coding
20	401.6	22.4	7688	20	AAO83834	Telomerase coding
21	401.6	22.4	7797	20	AAO83835	Telomerase coding
22	388.4	21.7	4222	10	AAO83836	Sequence of recomb
23	388.4	21.7	4466	10	AAO83837	Sequence of recomb
24	388.4	21.7	4466	15	AAO83838	Sequence of recomb
25	386.8	21.6	4207	10	AAO83839	Sequence of recomb
26	385.2	21.5	4204	12	AAO83840	Sequence of recomb
27	336	18.7	5641	22	AAO83841	Sequence of recomb
28	336	18.7	5670	22	AAO83842	Sequence of recomb
29	336	18.7	5826	22	AAO83843	Sequence of recomb
30	336	18.7	6071	22	AAO83844	Sequence of recomb
31	332.8	18.6	3753	24	AAO83845	Sequence of recomb
32	332.8	18.6	3753	24	AAO83846	Sequence of recomb
33	332.8	18.6	3950	22	AAO83847	Sequence of recomb
34	332.8	18.6	4176	10	AAO83848	Sequence of recomb
35	332.8	18.6	4214	21	AAO83849	Sequence of recomb
36	332.8	18.6	4214	21	AAO83850	Sequence of recomb
37	332.8	18.6	4357	22	AAO83851	Sequence of recomb
38	332.8	18.6	4816	22	AAO83852	Sequence of recomb
39	332.8	18.6	5021	21	AAO83853	Sequence of recomb
40	332.8	18.6	5027	20	AAO83854	Sequence of recomb
41	332.8	18.6	5069	20	AAO83855	Sequence of recomb
42	332.8	18.6	5069	20	AAO83856	Sequence of recomb
43	332.8	18.6	5099	20	AAO83857	Sequence of recomb
44	332.8	18.6	5104	22	AAO83858	Sequence of recomb
45	332.8	18.6	5132	20	AAO83859	Sequence of recomb

ALIGNMENTS

RESULT 1	AAA47190	standard: DNA: 1793 BP.
AAA47190	AAA47190	standard: DNA: 1793 BP.
AC	AAA47190:	
XX		
DT	03-OCT-2000	(first entry)
DE	Nucleotide sequence of the expression cassette of pRPA-BCAT41.	
XX		
KW	Methionine; 2-hydroxy-4-methylthiobutanoic acid; nitrilase;	
KW	nitrile hydratase; amidase; pRPA-BCAT41; ss.	
XX		
OS	Synthetic.	
XX		
FT	Key	location/Qualifiers
FT	CDS	123..1193
FT		/*tag= a
XX		
PN	WO200036120-A1.	
XX		
PD	22-JUN-2000.	
XX		
PF	10-DEC-1999;	99WO-FR03089.
XX		
PR	11-DEC-1998;	98FR-0015849.
PR	19-JUL-1999;	99FR-0009489.
XX		
PA	(RHON) RHONE-POULENC ANIMAL NUTRITION SA.	
XX		
PI	Favre-Bulle O, Pierrard J, Batisse Deblitte N.	
XX		

DR WPI: 2000-431598/37.
 DR P-PSDB: AAY93908.
 XX selecting sequences encoding enzymes involved in methionine synthesis,
 PT useful for hydrolysis of nitrile groups, by transforming methionine,
 PT auxotrophs and selection for growth
 XX
 PS Example 1: Page 27-29; 38pp: French.
 XX
 CC The specification describes a process for the selection and/or isolation
 CC of DNA sequences that encode enzymes involved in bioconversion of
 CC substrates to methionine or its derivatives such as
 CC 2-hydroxy-4-methylthiobutanoic acid. DNA fragments are cloned into
 CC a microbial expression vector and recombinant vectors used to transform
 CC a host that is auxotrophic for methionine (Met). The cells are cultured
 CC in medium containing an adequate amount of substrate and microbes able
 CC to grow on this medium are selected and/or isolated. DNA sequences
 CC involved in conversion of substrates are then isolated and/or identified.
 CC The method is used to identify DNA sequences encoding nitrilases, nitrile
 CC hydrolases or amidases. Nitrilases are useful in many synthetic process
 CC that require hydrolysis of nitrile groups, e.g. for production of the
 CC hydroxy analogue of Met. The present sequence is the nucleotide sequence
 CC of the expression cassette of PRPA-BCAT41.
 CC
 XX
 SO Sequence 1793 BP: 412 A; 527 C; 478 G; 376 T; 0 other;

Query Match 100.0%; Score 1793; DB 21; Length 1793;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattccctgtgacaattaatcgaactagtaactagtcgacgcttgctcagc 60
 DB 1 gaattccctgtgacaattaatcgaactagtaactagtcgacgcttgctcagc 60
 QY 61 tcgaactcgcgcgaagcttgccatcatcaatcaatttctcaaggaatacttac 120
 DB 61 tcgaactcgcgcgaagcttgccatcatcaatcaatttctcaaggaatacttac 120
 QY 121 atatacagaacaagaaatcgtccggcagccgctacagcgccgtctcccaactag 180
 DB 121 atatacagaacaagaaatcgtccggcagccgctacagcgccgtctcccaactag 180
 QY 181 atctggcaacgggtgtgtgataaaacattgagctgctgctcagcgccgagatgag 240
 DB 181 atctggcaacgggtgtgtgataaaacattgagctgctgctcagcgccgagatgag 240
 QY 241 gtgaactgacgtgttctgtgaaacctggtccgagctatccctccacgcttgctg 300
 DB 241 gtgaactgacgtgttctgtgaaacctggtccgagctatccctccacgcttgctg 300
 QY 301 ggcacacggcctgtgctgtaaaatacagtgccgctactatgccaactgcctcgtg 360
 DB 301 ggcacacggcctgtgctgtaaaatacagtgccgctactatgccaactgcctcgtg 360
 QY 361 acagtgcaagattcaacgcatgcccagggccgacgagacttgggtatattcatcg 420
 DB 361 acagtgcaagattcaacgcatgcccagggccgacgagacttgggtatattcatcg 420
 QY 421 tgggtatagcagcgacggcgagcagccttactctgagccaaatccctatagcagaca 480
 DB 421 tgggtatagcagcgacggcgagcagccttactctgagccaaatccctatagcagaca 480
 QY 481 agggcagatgctgtgctgctgcgcacaaactcaacactcaacatgttgaagcacgctg 540
 DB 481 agggcagatgctgtgctgctgcgcacaaactcaacactcaacatgttgaagcacgctg 540
 QY 541 ttggtgaaaggtatacgcgcagatctgactgtgtcgcacacgagcttgggcgctgctg 600
 DB 541 ttggtgaaaggtatacgcgcagatctgactgtgtcgcacacgagcttgggcgctgctg 600
 QY 601 cccctgtgctgtgaggacccctgtccctctgagcaagtagcgctgactcccaagcag 660
 DB 601 cccctgtgctgtgaggacccctgtccctctgagcaagtagcgctgactcccaagcag 660

DB 601 cccctgtgctgtgaggacccctgtccctctgagcaagtagcgctgactcccaagcag 660
 QY 661 aagcattcacattgcgcgcctgctccttcttcgtctgtaacaacagagcccaagcgc 720
 DB 661 aagcattcacattgcgcgcctgctccttcttcgtctgtaacaacagagcccaagcgc 720
 QY 721 tcagcgcaaggtgaaatgctgctcgcaaatctattcgtgtgaagggcagtgctta 780
 DB 721 tcagcgcaaggtgaaatgctgctcgcaaatctattcgtgtgaagggcagtgctta 780
 QY 781 ccattgcgcgaacagatgctgtaaccagaagacattgacatgctgggaagtgggaac 840
 DB 781 ccattgcgcgaacagatgctgtaaccagaagacattgacatgctgggaagtgggaac 840
 QY 841 acacagcctccctcgtcgtgaagtgaggcgagcttcataatgttttgcgcggaagcag 900
 DB 841 acacagcctccctcgtcgtgaagtgaggcgagcttcataatgttttgcgcggaagcag 900
 QY 901 gcaatctgctcctcctacccgcacagatgctgcaagggctgatacttgcgcgtgaaaca 960
 DB 901 gcaatctgctcctcctacccgcacagatgctgcaagggctgatacttgcgcgtgaaaca 960
 QY 961 tggagaatctgcttcgcgcgaagggatcaagaccctgtgggcactactccaacccg 1020
 DB 961 tggagaatctgcttcgcgcgaagggatcaagaccctgtgggcactactccaacccg 1020
 QY 1021 aggcaccccgctcgtgacttggaaccttgaggcaccgtgagccatgactcgggtgcatcca 1080
 DB 1021 aggcaccccgctcgtgacttggaaccttgaggcaccgtgagccatgactcgggtgcatcca 1080
 QY 1081 aaagctgtatccaggaagaagctcccgagccgacgttgcaaaataagcttgccgcccgtg 1140
 DB 1081 aaagctgtatccaggaagaagctcccgagccgacgttgcaaaataagcttgccgcccgtg 1140
 QY 1141 ccgtcagcgaactcagagctcgtacgatacgttactgtgcaagaacccgtctgaccccaaa 1200
 DB 1141 ccgtcagcgaactcagagctcgtacgatacgttactgtgcaagaacccgtctgaccccaaa 1200
 QY 1201 agatgacaagcccgagcaaacctgtccgggtcttgatctccttcgtcccgatccact 1260
 DB 1201 agatgacaagcccgagcaaacctgtccgggtcttgatctccttcgtcccgatccact 1260
 QY 1261 agttctagaagtgacacttgacgatacgaagctgggtcccaactgagcccatgcccgaact 1320
 DB 1261 agttctagaagtgacacttgacgatacgaagctgggtcccaactgagcccatgcccgaact 1320
 QY 1321 cagaagtgaacacgcgtagagccgactgtagtctggtggtctcccaatgagagtaggga 1380
 DB 1321 cagaagtgaacacgcgtagagccgactgtagtctggtggtctcccaatgagagtaggga 1380
 QY 1381 actgcagagcatcaataaacaagaaggtcagtcgaagaacttggtccttctgttattc 1440
 DB 1381 actgcagagcatcaataaacaagaaggtcagtcgaagaacttggtccttctgttattc 1440
 QY 1441 tgttgttttgcgttgaaagcttcctcgtgagttaggaacaatccgcgggaagcggattgaa 1500
 DB 1441 tgttgttttgcgttgaaagcttcctcgtgagttaggaacaatccgcgggaagcggattgaa 1500
 QY 1501 gttgcgaagcaaacggccggagaggttgccgagcagcgcgcacataaactgcgaagcat 1560
 DB 1501 gttgcgaagcaaacggccggagaggttgccgagcagcgcgcacataaactgcgaagcat 1560
 QY 1561 caaatlaagcagaagcactcctgacgataggtccttcttgcgtttctacaactcttcc 1620
 DB 1561 caaatlaagcagaagcactcctgacgataggtccttcttgcgtttctacaactcttcc 1620
 QY 1621 gtcgtcatatctacaagcactcccccacagatacgttaactagctcgtttttgcatc 1680
 DB 1621 gtcgtcatatctacaagcactcccccacagatacgttaactagctcgtttttgcatc 1680
 QY 1681 aggaagcagctatgaaacacactccttaaaacccctggaacacatttgacattgacaata 1740
 DB 1681 aggaagcagctatgaaacacactccttaaaacccctggaacacatttgacattgacaata 1740

Oy 1741 gctcagcattgtatgtgtccgaagacgaacaacattractcaatgccgcg 1793
Db 1741 gctcagcattgtatgtgtccgaagacgaacaacattractcaatgccgcg 1793

RESULT 2

AAZ45325
ID AAZ45325 standard; DNA; 1793 BP.

AAZ45325;

27-MAR-2000 (first entry)

Nucleotide sequence of an expression cassette encoding a nitrilase.

XX Tryptophan promoter; Prrp promoter; heterologous protein expression;
KW Escherichia coli W; industrial protein production; enzyme: nitrilase; ss.

XX Synthetic.
OS Alcaligenes faecalis.

XX Key Location/Qualifiers
FT CDS 123..1193
FT /tag= a
FT /product= "nitrilase"

XX M09964607-A1.

XX 16-DEC-1999.

XX 08-JUN-1999; 99MO-FR01343.

XX 10-JUN-1998; 98FR-0007474.

XX (RHON) RHONE-POULENC NUTRITION ANIMALE.

XX Pierard J, Guitton C, Favre-Bulle O;

XX WPI: 2000-097541/08.
XX P-PSDB: MAY34121.

XX Industrial production of heterologous proteins in Escherichia coli
PT strain W, particularly for expressing enzymes -

XX Example 1: Page 36-38; 52pp; French.

XX The present sequence represents an expression cassette comprising
CC the tryptophan promoter (Prrp promoter) and DNA encoding an Alcaligenes
CC faccalis ATCC8750 nitrilase (nitB). The nitrilase polynucleotide and the
CC promoter sequence were extracted from plasmid pRPA6BQr6 by restriction
CC digest. The Prrp promoter is used to control the expression of a
CC heterologous protein in an expression cassette which is used to modify
CC a strain of Escherichia coli W. The modified strain is then used for
CC industrial production of heterologous proteins. Specifically, the
CC promoter is used to control the expression of an Alcaligenes nitrilase
CC gene. The method is especially used to produce proteins of relatively
CC low value, preferably enzymes and specifically nitrilases.

XX Sequence 1793 BP; 412 A; 527 C; 478 G; 376 T; 0 other:

Query Match 100.0%; Score 1793; DB 21; Length 1793;

Best Local Similarity 100.0%; Pred. NO. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1793; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gaattccgtgtgacaataatcatcgactagtaactagtagcagctgtgcag 60
Db 1 gaattccgtgtgacaataatcatcgactagtaactagtagcagctgtgcag 60
Oy 61 tcgacctgcagcgaagctgtgggcatatcatcaaatgttatcttaagaaatactac 120
Db 61 tcgacctgcagcgaagctgtgggcatatcatcaaatgttatcttaagaaatactac 120

Oy 121 atatgcagacagaanaaaatcgtccgggacgcgcgtacagcgcgcctctccaaatag 180
Db 121 atatgcagacagaanaaaatcgtccgggacgcgcgtacagcgcgcctctccaaatag 180
Oy 181 atctggcaacgggtgtgtataaaacatctgagctgtgctcgaagcccgcatagagct 240
Db 181 atctggcaacgggtgtgtataaaacatctgagctgtgctcgaagcccgcatagagct 240
Oy 241 gtgacctgacgtgtgttgtaaaacctgtgcccgtctaccccttccagctgtgctg 300
Db 241 gtgacctgacgtgtgttgtaaaacctgtgcccgtctaccccttccagctgtgctg 300
Oy 301 ggcgacccgctgtgtcgtctgtaataatagtcgccgttactatgccaactcgtctcgc 360
Db 301 ggcgacccgctgtgtcgtctgtaataatagtcgccgttactatgccaactcgtctcgc 360
Oy 361 acagtgcagagtttcaaacgattgccagcgacgacgagaccttggtaatttcatcgac 420
Db 361 acagtgcagagtttcaaacgattgccagcgacgacgagaccttggtaatttcatcgac 420
Oy 421 tgggtatagcagcagcgacgagcgagcgttaccttggccaatgctctgatcgacga 480
Db 421 tgggtatagcagcagcgacgagcgagcgttaccttggccaatgctctgatcgacga 480
Oy 481 agggcagatctgtgtgcgcgttcgcaaacctcaaacctacatgtttagcgcacagctgt 540
Db 481 agggcagatctgtgtgcgcgttcgcaaacctcaaacctacatgtttagcgcacagctgt 540
Oy 541 ttggtgaaggtatagtcgagacatctgattgttcggaacacgagcttggcgcgtctg 600
Db 541 ttggtgaaggtatagtcgagacatctgattgttcggaacacgagcttggcgcgtctg 600
Oy 601 cccgtgtcgtctgtggagacacctgtcccttgagcgaagtacgcgtctactccagcag 660
Db 601 cccgtgtcgtctgtggagacacctgtcccttgagcgaagtacgcgtctactccagcag 660
Oy 661 aagccattcacatgtccgcgcgttcgtcttctgtctgtacgagaaagcccatgctgc 720
Db 661 aagccattcacatgtccgcgcgttcgtcttctgtctgtacgagaaagcccatgctgc 720
Oy 721 tcaagcgaaggtatgacatgtgctcctgcgaatctatcgtgtgaagcgagctgtta 780
Db 721 tcaagcgaaggtatgacatgtgctcctgcgaatctatcgtgtgaagcgagctgtta 780
Oy 781 ccattgcgcgcaagcagttgtctcaaccagagacacttgacatgctgtaagtgtgac 840
Db 781 ccattgcgcgcaagcagttgtctcaaccagagacacttgacatgctgtaagtgtgac 840
Oy 841 acaagcctccctgtctgaagtgtggcgcgagcttcacatgttttgcgcgcgagac 900
Db 841 acaagcctccctgtctgaagtgtggcgcgagcttcacatgttttgcgcgcgagac 900
Oy 901 gacatgtgtctccactacctgcacacgaltgcgaagcgctgtcatctgcgactgaaca 960
Db 901 gacatgtgtctccactacctgcacacgaltgcgaagcgctgtcatctgcgactgaaca 960
Oy 961 tggagaagaattgtccttcgcaagcgatcaaacgacctgtgtggcactactcaaaccc 1020
Db 961 tggagaagaattgtccttcgcaagcgatcaaacgacctgtgtggcactactcaaaccc 1020
Oy 1021 aggcacccgctctgtactgtgaccttgggcaacgagcccatctgactcgtggtatcca 1080
Db 1021 aggcacccgctctgtactgtgaccttgggcaacgagcccatctgactcgtggtatcca 1080
Oy 1081 aaagcgtatccagaagaagctcccgagcgacgctgcaaatgacgctgcgcgcgtcg 1140
Db 1081 aaagcgtatccagaagaagctcccgagcgacgctgcaaatgacgctgcgcgcgtcg 1140
Oy 1141 ccgtacgcagactcagagactggtacgtactgtgtgcaagaacgctcgtgaaccccaa 1200
Db 1141 ccgtacgcagactcagagactggtacgtactgtgtgcaagaacgctcgtgaaccccaa 1200

QY 1201 agatgacaaagcccgagcaactgtccgggtcttgatctctcttggtggtcccgatccact 1260
CC |
CC |
DB 1201 agatgacaaagcccgagcaactgtccgggtcttgatctctcttggtggtcccgatccact 1260
QY 1261 agtctcgaagtcgacccgagcaatgcaagcttggtgtccaccctgaccctatcccgact 1320
CC |
CC |
DB 1261 agtctcgaagtcgacccgagcaatgcaagcttggtgtccaccctgaccctatcccgact 1320
QY 1321 cagaagtgaacacgcccgtacgacgtagtctggtgtctccaccatgacgaagtaagga 1380
CC |
CC |
DB 1321 cagaagtgaacacgcccgtacgacgtagtctggtgtctccaccatgacgaagtaagga 1380
QY 1381 actgcccagatcaataataaagcctcagtcgaaagactgagcccttcgtttatc 1440
DB 1381 actgcccagatcaataataaagcctcagtcgaaagactgagcccttcgtttatc 1440
QY 1441 tgtgtgtgtcgttgaaagctctccctgagtaggacaataccgcccggagcgattgaa 1500
CC |
CC |
DB 1441 tgtgtgtgtcgttgaaagctctccctgagtaggacaataccgcccggagcgattgaa 1500
QY 1501 gtctcgaaagcaacgcccggaggtggtggtgacgagacgcccgcataaactgacagat 1560
DB 1501 gtctcgaaagcaacgcccggaggtggtggtgacgagacgcccgcataaactgacagat 1560
QY 1561 caaatgaagcaagaagccatctcgtgacgagatgaccttctgctttctacaactctcct 1620
DB 1561 caaatgaagcaagaagccatctcgtgacgagatgaccttctgctttctacaactctcct 1620
QY 1621 gtctgcatcatcaaaagcctcccccacagatagcgttaactagacctgttttgatc 1680
DB 1621 gtctgcatcatcaaaagcctcccccacagatagcgttaactagacctgttttgatc 1680
QY 1681 aggaagcagcgtatgaaccactccttaaaaacccgtgaaacattgtgcatgtacataat 1740
DB 1681 aggaagcagcgtatgaaccactccttaaaaacccgtgaaacattgtgcatgtacataat 1740
QY 1741 gctcagacactgtgtgtgtcgcgaagacgaacaatactcaatgccgcg 1793
DB 1741 gctcagacactgtgtgtgtcgcgaagacgaacaatactcaatgccgcg 1793

RESULT 3
AA086279
ID AA086279 standard; DNA: 1727 BP.
XX
AC AA086279;
XX
DT 12-AUG-1999 (first entry)
XX
DE Nitrilase enzyme coding sequence.
XX
KW Nitrilase: organic acid preparation; nitrile compound hydrolysis;
KW amide production; carboxylic acid production; ss.
XX
OS Alcaligenes faecalis.
XX
PN JP06153968-A.
XX
PD 03-JUN-1994.
XX
PF 17-NOV-1992: 92JP-0306663.
XX
PR 17-NOV-1992: 92JP-0306663.
XX
PA (NITR) NITRO CHEM IND CO LTD.
XX
DR WPI: 1994-220515/27.
XX
DR P-PSDB: AAR61104.
XX
PT Organic acid prepn by hydrolysing nitrile - using nitrilase enzyme
XX obtd from transformed organism.
PS Claim 1; Fig 2; 9pp; Japanese.

XX This sequence encodes the nitrilase enzyme from *Alcaligenes faecalis*.
CC The invention relates to a method of preparing organic acid by
CC hydrolysing nitrile compounds of formula R1CH(R2)CN with nitrilase to
CC form carboxylic acid of formula R1CH(R2)COOH and equimolar amount of
CC ammonia. (R1 is optionally substituted phenyl, naphthyl, indole,
CC thiophene, pyridyl or cyclohexenyl; R2 is H or hydroxyl group). The
CC nitrilase is prepared by culturing a transformant containing recombinant
CC DNA obtained by integrating into a vector the A. faecalis nitrilase
CC enzyme DNA. Various kinds of amide or carboxylic acid can be produced
CC industrially in good efficiency.
SQ Sequence 1727 BP; 343 A; 566 C; 468 G; 350 T; 0 other:

Query Match 62.8%; Score 1126.4; DB 15; Length 1727;
Best Local Similarity 99.5%; Pred. No. 3,5e-234;
Matches 1130; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 122 tatgacagaagaanaaaatcgctccggagcgcgcgtacagccgcgtctcccaactacga 181
DB 122 tatgacagaagaanaaaatcgctccggagcgcgcgtacagccgcgtctcccaactacga 181
QY 182 tctggcaacgggtgttgataaaacattgagctgctcgtcagcccgacgaatgaagctg 241
DB 182 tctggcaacgggtgttgataaaacattgagctgctcgtcagcccgacgaatgaagctg 241
QY 242 tgaacctatcgtctgttgtaaaccttggtccgcgcgtacacctccacgtctgctggg 301
DB 242 tgaacctatcgtctgttgtaaaccttggtccgcgcgtacacctccacgtctgctggg 301
QY 302 cgcacccgctgtgctcgtgaataacagtgccgactactatgacaaactgctcgtcgtga 361
DB 302 cgcacccgctgtgctcgtgaataacagtgccgactactatgacaaactgctcgtcgtga 361
QY 362 cagtgcgaagttccaagcattgcccagcgccgacgacctgggtatttcatctgcac 421
DB 362 cagtgcgaagttccaagcattgcccagcgccgacgacctgggtatttcatctgcac 421
QY 421 cagtcgacgagttccaagcattgcccagcgccgacgacctgggtatttcatctgcac 602
DB 421 cagtcgacgagttccaagcattgcccagcgccgacgacctgggtatttcatctgcac 602
QY 422 ggtttatagcagcgacgacgacgaccttacccttgagcgaactgctgaatcgacgaaca 481
DB 422 ggtttatagcagcgacgacgacgaccttacccttgagcgaactgctgaatcgacgaaca 481
QY 481 gggccagatgctgtgtcgcgctcgcgaactcaaaactacaacatgttgacgacgctgt 541
DB 481 gggccagatgctgtgtcgcgctcgcgaactcaaaactacaacatgttgacgacgctgt 541
QY 542 tggtagaaggtatgcccagatctgtatgtgtcgcgaacacggaactggccgctgggtg 601
DB 542 tggtagaaggtatgcccagatctgtatgtgtcgcgaacacggaactggccgctgggtg 601
QY 602 cctgtgtcgttggaagacactgtcccccctgagcaagtaacgctgttaccacgaacga 661
DB 602 cctgtgtcgttggaagacactgtcccccctgagcaagtaacgctgttaccacgaacga 661
QY 662 agccatcacatctgcgcgtcgtccctcttctgctgtacagcgaaagcgccatgcgt 721
DB 662 agccatcacatctgcgcgtcgtccctcttctgctgtacagcgaaagcgccatgcgt 721
QY 722 cagcgccaaggtgaacatgctcctcgcgaactatcttggttgaagcgacgtcttac 781
DB 722 cagcgccaaggtgaacatgctcctcgcgaactatcttggttgaagcgacgtcttac 781
QY 782 catcgccgacgacgtgtgtctaccccgaagacacgtgacatcttggaagtgaagtaaga 841
DB 782 catcgccgacgacgtgtgtgtctaccccgaagacacgtgacatcttggaagtgaagtaaga 841
QY 842 caacgctctcctgtcgaagaatggcgcgacgttccatgatttctgcgcggaacggaag 901
DB 842 caacgctctcctgtcgaagaatggcgcgacgttccatgatttctgcgcggaacggaag 901
QY 902 cacatggtcctcctacgtccacacagatgcccgaagcgtgatatgcccgatgtgaacat 961
DB 902 cacatggtcctcctacgtccacacagatgcccgaagcgtgatatgcccgatgtgaacat 961

Db 1083 cacatlggtccctactcgtgccacagatgcccgaagcgctgatalcattgcccgaatctgaacat 1142
Qy 962 ggaagaattcctcttcgcaagcgatcaacgacctgtggccaactctccaacccga 1021
Db 1143 ggaagaattcctcttcgcaagcgatcaacgacctgtggccaactctccaacccga 1202
Qy 1022 ggcacaccgtctgtactgagacgttgagcaccgttgagcccttgatctgggtgcatccaa 1081
Db 1203 ggcacaccgtctgtactgagacgttgagcaccgttgagcccttgatctgggtgcatccaa 1262
Qy 1082 aagcgtgatacgaagaagactcccgagccgacgttgcaaaagtacgctgcgcgcgtgc 1141
Db 1263 aagcgtgatacgaagaagactcccgagccgacgttgcaaaagtacgctgcgcgcgtgc 1322
Qy 1142 cgtacagcagactcaagactcgatagctactggtgtaagaagaacctctgaccccaaa 1201
Db 1323 cgtacagcagactcaagactcgatagctactggtgtaagaagaacctctgaccccaaa 1382
Qy 1202 gatgacaagcgccgcaaacctgtccggctcttgatctctctcgtcccgatcc 1257
Db 1383 gatgacaagcgccgcaaacctgtccggctcttgatctctctcgtcccgatcc 1438

RESULT 4

AA12173

ID AA12173 standard; DNA; 1071 BP.

XX

AC AA12173;

DT 10-AUG-2000 (first entry)

XX

DE A. faecalis nitrilase DNA.

XX

KW Nitrilase; chiral carboxylic acid; racemic nitrile; pharmaceutical;

KM plant protection agent; ds.

XX

OS Alcaligenes faecalis.

XX

PN DEJ9848129-A1.

XX

PD 20-APR-2000.

XX

PF 19-OCT-1998; 98DE-1048129.

XX

PR 19-OCT-1998; 98DE-1048129.

XX

PA (BADT) BASF AG.

XX

PI Rees-Loeschke M, Friedrich T, Hauer B, Matthes R, Engels D;

XX

DR WPI: 2000-319322/28.

XX

DR P-PSDB: MAY87649.

XX

PT New nucleic acid sequence encoding Alcaligenes faecalis nitrilase

XX

PT polypeptide useful for converting racemic nitriles to chiral carboxylic

XX

PS Claim 1a; Page 14-16; 28pp; German.

XX

CC This invention describes a novel nucleic acid sequence (I) encoding an

XX

CC Alkaligenes faecalis nitrilase polypeptide. The polypeptide encoded by

XX

CC (I), or a growing, resting or digested microorganism containing (I) is

XX

CC useful for preparing chiral carboxylic acids (III) by conversion of

XX

CC racemic nitriles (IV). (III) are useful as intermediates, e.g. for

XX

CC pharmaceuticals and plant protection agents. This sequence encodes the

XX

CC Alkaligenes faecalis nitrilase described in the method of the invention.

XX

Sequence 1071 BP; 230 A; 318 C; 304 G; 219 T; 0 other;

Query Match 53.28; Score 954.2; DB 21; Length 1071;
Best Local Similarity 93.28; Pred. No. 5.4e-197;
Matches 998; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Qy 123 atgcagacaagaaaaatcgtccggcagcgccgtatcacagcgccgtctcccaactacgat 182
Db 1 atgcagacaagaaaaatcgtccggcagcgccgtatcacagcgccgtctcccaactacgat 60
Qy 183 ctgccaacgggtgtgtgataaacaatctagctgtgtctgcagcccgagatgaaggcgtgt 242
Db 61 ctgccaacgggtgtgtgataaacaatctagctgtgtctgcagcccgagatgaaggcgtgt 120
Qy 243 gaactgatcgtgtgtgtgataaactgtgtcccggtctatccctccacgctgtgtgtgtgc 302
Db 121 gaactgatcgtgtgtgtgataaactgtgtcccggtatctccctccacgctgtgtgtgc 180
Qy 303 gcaacggcctgtgtgtgataaatacagtgcccggtactatgtcaactgtgtgtgtgtgtgc 362
Db 181 gcaacggcctgtgtgtgataaatacagtgcccggtactatgtcaactgtgtgtgtgtgtgc 240
Qy 363 agtgaagattcaacgcatctgcccagcgccgacgagacttgggtatttcatcgcactgt 422
Db 241 agtgaagattcaacgcatctgcccagcgccgacgagacttgggtatttcatcgcactgt 300
Qy 423 ggtataagcagcgacgagcgccgacgacttcaactgtggccaatgtcgtatcgaagacaag 482
Db 301 ggtataagcagcgacgagcgccgacgacttcaactgtggccaatgtcgtatcgaagacaag 360
Qy 483 ggcacagatcgt 542
Db 361 ggcacagatcgt 420
Qy 543 ggtgaaggtatgtcccgagatctgattgtgtcgcgacacacgagctgtgagcggtcgtgtcc 602
Db 421 ggtgaaggtatgtcccgagatctgattgtgtcgcgacacacgagctgtgagcggtcgtgtcc 480
Qy 603 ctgt 662
Db 481 ctgt 540
Qy 663 ggcattacattgcccgcctgtgcccgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 722
Db 541 ggcattacattgcccgcctgtgcccgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Qy 723 agcgccaaggtgaacatgt 782
Db 601 agtgccaaggtgaacatgt 660
Qy 783 atgcgcgcagcagatgt 842
Db 661 atgcgcgcagcagatgt 720
Qy 843 aagcgtccctgtgataaggt 902
Db 721 aagcgtccctgtgataaggt 780
Qy 903 aactgtgtccctactgtgccacagatgtccgaaagcgctgatactgtccgatactgtacatgt 962
Db 781 aactgtgtccctactgtgccacagatgtccgaaagcgctgatactgtccgatactgtacatgt 840
Qy 963 gaagaattcctctgcgacaagcgatcaacgacccctgtgtgtgtgtgtgtgtgtgtgtgtgt 1022
Db 841 gaagaattcctctgcgacaagcgatcaacgacccctgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Qy 1023 gccacccgtctgt 1082
Db 901 gccacccgtctgt 960
Qy 1083 agcgtgtatccaggaagaagctcccgagccgacgagtgtaaaagtacgctgtgagccgtgtcc 1142
Db 961 agcgtgtatccaggaagaagctcccgagccgacgagtgtaaaagtacgctgtgagccgtgtcc 1020
Qy 1143 gtcagcagactcaagactcgatacgtactgtgtgcaagaacgctgtccta 1193
Db 1021 atcagccatccacagactcgacacactgtgactagtgcaagaagcgctgtga 1071

```

RESULT      5
AA083824
ID   AA083824 standard; DNA; 5096 BP.
XX
AC   AA083824;
XX
DT   04-OCT-1995 (first entry)
XX
DE   Plasmid pKK232-8.
XX
KM   Plasmid pKK232-8; bacteriophage T7; RNA polymerase;
XX   mutant promoters; circular; cyclic; ds.
XX
OS   Synthetic.
XX
FH   Key
FT   misc-feature
FT   /tag=a
FT   /note="Cloning polylinker prior to a promoterless
XX   chloramphenicol acetyl transferase gene"
XX
PN   US5385834-A.
XX
PD   31-JAN-1995.
XX
PF   13-AUG-1993; 93US-0106433.
XX
PR   13-AUG-1993; 93US-0106433.
XX
PA   (GEOR-) GEORGIA TECH RES CORP.
XX
PI   Ikeda RA;
XX
DR   WPI; 1995-081565/11.
XX
PT   New mutant form of T7 RNA polymerase - able to utilise a wide
PT   range of T7 promoter sequences having mutations that inactivate
PT   recognition by wild type enzyme
XX
PS   Disclosure; Columns 37-42; 79pp; English.
XX
CC   AA083824 is the plasmid pKK232-8, it was used in the production of
CC   claimed bacteriophage T7 RNA polymerase mutant promoters.
XX
SQ   Sequence 5096 BP; 1294 A; 1290 C; 1290 G; 1222 T; 0 other;

Query Match      22.5%; Score 404; DB 16; Length 5096;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 agagtcgacccgagcagcatgcaagcttggtgccacgtgaccccatgcccgaactcagaag 1326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1237 agaatctgcttgccgagcagtagcggtggtccacactgaccccatgcccgaactcagaag 1296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1327 tgaacgcgctagcgcgcgagtagtctgtgggtctcccatgcccgaagtagggaaactgcc 1386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1297 tgaacgcgctagcgcgcgagtagtctgtgggtctcccatgcccgaagtagggaaactgcc 1356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1387 aggcacataaataaagaagaagctcagtcgaagaagactgggaccttgcgtttatctgtgtg 1446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1357 aggcacataaataaagaagaagctcagtcgaagaagactgggaccttgcgtttatctgtgtg 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1447 ctgtcgttgaacgctctccctgagtagagacaatacgcgcgggaacggaattggaacgttgcg 1506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1417 ctgtcgttgaacgctctccctgagtagagacaatacgcgcgggaacggaattggaacgttgcg 1476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1507 aagcaacgcgcgcgaggtgtgctggcgaggaagcgcgcgcataaactgtccaggaatcaaat 1566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1477 aagcaacgcgcgcgaggtgtgctggcgaggaagcgcgcgcataaactgtccaggaatcaaat 1536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1567 aagcaaaaggcctactctgaagatgaccttcttgcgtttctacaactctctctgtcgtc 1626

```

```

DB 1537 aagcaaaaggcctactctgaagatgaccttcttgcgtttctacaactctctctgtcgtc 1596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1627 atactcaagccatccccccacagatacggtaactagacctgttttgcatcagaaga 1686
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1597 atactcaagccatccccccacagatacggtaactagacctgttttgcatcagaaga 1656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1687 gcagctat 1694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1657 gcagctctg 1664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT      6
AA083833
ID   AA083833 standard; DNA; 5110 BP.
XX
AC   AA083833;
XX
DT   05-OCT-1995 (first entry)
XX
DE   Plasmid PCM-B64.
XX
KM   Plasmid PCM-B64; bacteriophage T7; RNA polymerase;
XX   mutant -8T to C; T7 promoter phi 10; circular; cyclic; ds.
XX
OS   Synthetic.
XX
FH   Key
FT   Promoter
FT   /tag=a
FT   /note="Intermediate T7 promoter phi 10 mutant (-8T
XX   to C) prior to a promoterless chloramphenicol
XX   acetyl transferase gene"
XX
PN   US5385834-A.
XX
PD   31-JAN-1995.
XX
PF   13-AUG-1993; 93US-0106433.
XX
PR   13-AUG-1993; 93US-0106433.
XX
PA   (GEOR-) GEORGIA TECH RES CORP.
XX
PI   Ikeda RA;
XX
DR   WPI; 1995-081565/11.
XX
PT   New mutant form of T7 RNA polymerase - able to utilise a wide
PT   range of T7 promoter sequences having mutations that inactivate
PT   recognition by wild type enzyme
XX
PS   Disclosure; Columns 97-104; 79pp; English.
XX
CC   AA083833 is the plasmid PCM-B64, it was used in the production of
CC   the bacteriophage T7 RNA polymerase mutant promoter phi 10 -8T
CC   to C substitution (claimed).
XX
SQ   Sequence 5110 BP; 1305 A; 1289 C; 1289 G; 1227 T; 0 other;

Query Match      22.5%; Score 404; DB 16; Length 5110;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 agagtcgacccgagcagcatgcaagcttggtgccacgtgaccccatgcccgaactcagaag 1326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1251 agaatctgcttgccgagcagtagcggtggtccacactgaccccatgcccgaactcagaag 1310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1327 tgaacgcgctagcgcgcgagtagtctgtgggtctcccatgcccgaagtagggaaactgcc 1386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1311 tgaacgcgctagcgcgcgagtagtctgtgggtctcccatgcccgaagtagggaaactgcc 1370

```

[illegible]

RESULT	7
AA083834	
ID	AA083834 standard; DNA: 5110 BP.
XX	
AC	AA083834;
XX	
DT	05-OCT-1995 (first entry)
XX	
DE	Plasmid PCM-P1208.
XX	
KW	Plasmid PCM-P1208; bacteriophage T7; RNA polymerase;
XX	mutant -7C to G; T7 promoter phi 10; circular; cyclic; ds.
OS	Synthetic.
XX	
Key	Location/Qualifiers
FT	Promoter
FT	198..220
FT	/"tag="a
FT	/note="inactive T7 promoter phi 10 mutant (-7C to
FT	g) prior to a promoterless chloramphenicol
FT	acetyl transferase gene"
XX	
PN	US5385834-A.
XX	
PD	31-JAN-1995.
XX	
PE	13-AUG-1993; 93US-0106433.
XX	
PR	13-AUG-1993; 93US-0106433.
XX	
RA	(GEOR-) GEORGIA TECH RES CORP.
XX	
RA	Ikeda RA:
P1	
XX	
XX	WPI: 1995-081565/11.
XX	
PT	New mutant form of T7 RNA polymerase - able to utilise a wide
PT	range of T7 promoter sequences having mutations that inactivate
PT	recognition by wild type enzyme
XX	
PS	Disclosure; Columns 105-110; 79pp; English.
XX	
CC	AA083834 is the plasmid PCM-P1208, it was used in the production of
CC	the bacteriophage T7 RNA polymerase mutant promoter phi 10 -7C
CC	to G substitution (claimed).
XX	
SO	Sequence 5110 BP; 1305 A; 1290 G; 1227 T; 0 other;

Query Match	22.5%;	Score 404;	DB 16;	Length 5110;
Best Local Similarity	96.5%;	Pred. No. 4.9e-78;		
Matches 413;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;

QY	1267	aggtcgacgttcgacgcatgcaagcgttgggtcccaactggaacccaatgcccgaacccagaag	1326
Dp	1251	aggaattcgccctgcgcgcatagcgcgcgtgggtccccaactggaacccaatgcgaactcagaag	1310
QY	1327	tgaacgcgcgttaagcgcgcgaatgtagtgtgggtgtctcccatgtagaagtgaaggaaatgccc	1386
Dp	1311	tgaacgcgcgttaagcgcgcgaatgtagtgtgggtgtctcccatgtagaagtgaaggaaatgccc	1370
QY	1387	agggatcaaatataaaggaaggtctggtctgaaagactggagccttcggtttatctgtgt	1446
Dp	1371	agggatcaaatataaaggaaggtctggtctgaaagactggagccttcggtttatctgtgt	1430
QY	1447	tgtgcgtgtgaacgactctccctgaatgaagacaataatccgcgcggagcgcgaattgaacgttgcg	1506
Dp	1431	tgtgcgtgtgaacgactctccctgaatgaagacaataatccgcgcggagcgcgaattgaacgttgcg	1490
QY	1507	aagcaacgcgcgcgcgcgaggtgtggtgcgcgcgacgagacgcgcgaataaactgcgaagcatcaaat	1566
Dp	1491	aagcaacgcgcgcgcgcgaggtgtggtgcgcgcgacgagacgcgcgaataaactgcgaagcatcaaat	1550
QY	1567	aagcagaagaagccatccctcgaagaagatggtccttttgcgctttctcaaaccttcctcgtctgc	1626
Dp	1551	aagcagaagaagccatccctcgaagaagatggtccttttgcgctttctcaaaccttcctcgtctgc	1610
QY	1627	atactcaagaagcatccccccacagatacagtaacatgacctgcttttgcatacagaana	1686
Dp	1611	atactcaagaagcatccccccacagatacagtaacatgacctgcttttgcatacagaana	1670
QY	1687	gcagctat	1694
Dp	1671	gcagctgt	1678

Result	ID	Location/Qualifiers
AAQ83835	AAQ83835	standard; DNA; 5110 BP.
AAQ83835	AAQ83835	
06-OCT-1995	(first entry)	
Plasmid	PCM-P1031	
mutant PCM-P1031	bacteriophage T7; RNA polymerase;	
mutant -7C to A; T7 promoter	phi 10; circular; cyclic; ds.	
Synthetic.		
Key		
Promoter	198..220	
FT	/*tag=	
FT	a	
FT	/note= "Inactive T7 promoter phi 10 mutant (-7C to	
FT	A) prior to a promoterless chloramphenicol	
FT	acetyl transferase gene"	
US5385834-A.		
31-JAN-1995.		
13-AUG-1993;	930S-0106433.	
13-AUG-1993;	930S-0106433.	
(GEOR-)	GEORGIA TECH RES CORP.	
Ikeda RA;		
WPI; 1995-081565/11.		

```
XX New mutant form of T7 RNA polymerase - able to utilise a wide
PT range of T7 promoter sequences having mutations that inactivate
PT recognition by wild type enzyme
XX
XX Disclosure: Columns 111-118: 79pp; English.
XX
XX AA083835 is the plasmid pCM-T1031, it was used in the production of
CC the bacteriophage T7 RNA polymerase mutant promoter phi 10 -7C
CC to A substitution (claimed).
XX
SQ Sequence 5110 BP; 1306 A; 1286 C; 1290 G; 1228 T; 0 other;

Query Match      22.5%; Score 404; DB 16; Length 5110;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 agagtcgacctgacgacatgcaagcttggtcccaactgaccccatgccaactcagaag 1326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1251 agaatctgcctgcygcagtagcgcggtggtcccaactgaccccatgccaactcagaag 1310
QY 1327 tgaacgcgcgtagcgccgagtagtggtggtgtcccaactgaccccatgccaactcaga 1386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1311 tgaacgcgcgtagcgccgagtagtggtggtgtcccaactgaccccatgccaactcaga 1370
QY 1387 aggcatacaataaagcaagagctcagtcgaagaactgagccttcgtttatctgtgt 1446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1371 aggcatacaataaagcaagagctcagtcgaagaactgagccttcgtttatctgtgt 1430
QY 1447 ttgtcgttgaagcgtctctcgtgtagtagacaatccgcgcgagcgagattgaaactgtcg 1506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1431 ttgtcgttgaagcgtctctcgtgtagtagacaatccgcgcgagcgagattgaaactgtcg 1490
QY 1507 aagcaacgcccggaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1491 aagcaacgcccggaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1550
QY 1567 aagcaagaagcgcattcctgacgtagtggtccttttggttttctacaactcttctgtctgc 1626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1551 aagcaagaagcgcattcctgacgtagtggtccttttggttttctacaactcttctgtctgc 1610
QY 1627 atatctacaagcattcccccacagatacggttaactagcctgttttgcataaggaaa 1686
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1611 atatctacaagcattcccccacagatacggttaactagcctgttttgcataaggaaa 1670
QY 1687 gcagctat 1694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1671 gcagctgt 1678

RESULT 9
AA083836
ID AA083836 standard; DNA: 5110 BP.
XX
XX AA083836;
AC
XX 06-OCT-1995 (first entry)
DT
XX
XX Plasmid pCM-T221.
DE
XX
XX Plasmid pCM-T221; bacteriophage T7; RNA polymerase;
KW mutant -6A to G; T7 promoter phi 10; circular; cyclic; ds.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Promoter 198..220
FT /tag= a
FT /note= "Intermediate T7 promoter phi 10 mutant (-6A
FT to G) prior to a promoterless chloramphenicol
XX acetyl transferase gene"
```

```
PN 055385834-A.
XX
XX 31-JAN-1995.
PD
XX
XX 13-AUG-1993; 9305-0106433.
PF
XX
XX 13-AUG-1993; 9305-0106433.
PR
XX
XX 13-AUG-1993; 9305-0106433.
PA
XX
XX (GEOR-) GEORGIA TECH RES CORP.
XX
XX Ikeda RA.
PI
XX
XX WPI; 1995-081565/11.
XX
XX
XX New mutant form of T7 RNA polymerase - able to utilise a wide
PT range of T7 promoter sequences having mutations that inactivate
PT recognition by wild type enzyme
XX
XX Disclosure: Columns 119-124: 79pp; English.
XX
XX AA083836 is the plasmid pCM-T221, it was used in the production of
CC the bacteriophage T7 RNA polymerase mutant promoter phi 10 -6A
CC to G substitution (claimed).
XX
SQ Sequence 5110 BP; 1304 A; 1288 C; 1290 G; 1228 T; 0 other;

Query Match      22.5%; Score 404; DB 16; Length 5110;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1267 agagtcgacctgacgacatgcaagcttggtcccaactgaccccatgccaactcagaag 1326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1251 agaatctgcctgcygcagtagcgcggtggtcccaactgaccccatgccaactcagaag 1310
QY 1327 tgaacgcgcgtagcgccgagtagtggtggtgtcccaactgaccccatgccaactcaga 1386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1311 tgaacgcgcgtagcgccgagtagtggtggtgtcccaactgaccccatgccaactcaga 1370
QY 1387 aggcatacaataaagcaagagctcagtcgaagaactgagccttcgtttatctgtgt 1446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1371 aggcatacaataaagcaagagctcagtcgaagaactgagccttcgtttatctgtgt 1430
QY 1447 ttgtcgttgaagcgtctctcgtgtagtagacaatccgcgcgagcgagattgaaactgtcg 1506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1431 ttgtcgttgaagcgtctctcgtgtagtagacaatccgcgcgagcgagattgaaactgtcg 1490
QY 1507 aagcaacgcccggaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1491 aagcaacgcccggaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1550
QY 1567 aagcaagaagcgcattcctgacgtagtggtccttttggttttctacaactcttctgtctgc 1626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1551 aagcaagaagcgcattcctgacgtagtggtccttttggttttctacaactcttctgtctgc 1610
QY 1627 atatctacaagcattcccccacagatacggttaactagcctgttttgcataaggaaa 1686
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1611 atatctacaagcattcccccacagatacggttaactagcctgttttgcataaggaaa 1670
QY 1687 gcagctat 1694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1671 gcagctgt 1678

RESULT 10
AA083825
ID AA083825 standard; DNA: 5110 BP.
XX
XX AA083825;
AC
XX 04-OCT-1995 (first entry)
DT
XX
XX Plasmid pCM-X#.
```



```

XX XX Plasmid pCM-X#: bacteriophage T7; RNA polymerase;
KM mutant promoters: circular; cyclic; ds.
XX OS Synthetic.
XX Key Location/Qualifiers
FH Promoter 198..220
FT /tag= a
FT /note= "potential T7 promoter prior to a promoterless
chloramphenicol acetyl transferase gene"
XX US5385834-A.
XX 31-JAN-1995.
XX 13-AUG-1993; 93US-0106433.
XX 13-AUG-1993; 93US-0106433.
XX (GBOR-) GEORGIA TECH RES CORP.
XX Ikeda RA:
XX WPI: 1995-081565/11.
XX New mutant form of T7 RNA polymerase - able to utilise a wide
PT range of T7 promoter sequences having mutations that inactivate
PT recognition by wild type enzyme
XX PS Disclosure: Columns 43-50; 79pp; English.
XX AA083825 is the plasmid pCM-X#, it was used in the production of
CC claimed bacteriophage T7 RNA polymerase mutant promoters.
XX CC Sequence 5110 BP; 1303 A; 1287 C; 1288 G; 1226 T; 6 other;
SQ
Query Match 22.5%; Score 404; DB 16; Length 5110;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY 1267 agagtcgacctgcagcagctgtggtccacactgacccatgcccgaactcagaag 1326
DB 1251 agaatttcctgcgagcagtagcgcggtgtccacactgacccatgcccgaactcagaag 1310
OY 1327 tgaacgcgctagcgcgagtagtgtgtggtgtcccccagtagaggaactgccc 1386
DB 1311 tgaacgcgctagcgcgagtagtgtgtggtgtcccccagtagaggaactgccc 1370
OY 1387 aggcataataaagcgaagctcagtcgaaagactggtgcttcttcttctgtgt 1446
DB 1371 aggcataataaagcgaagctcagtcgaaagactggtgcttcttcttctgtgt 1430
OY 1447 ttgtcgtggaacgctctctctgtagtagacaatcccgccggagcgagattggaactgtgcg 1506
DB 1431 ttgtcgtggaacgctctctctgtagtagacaatcccgccggagcgagattggaactgtgcg 1490
OY 1507 aagcaacgcccggaggtgtgcggaagcgcgccgcataaactgccaagcatcaatt 1566
DB 1491 aagcaacgcccggaggtgtgcggaagcgcgccgcataaactgccaagcatcaatt 1550
OY 1567 aagcgaagcgcatctctgtagtagtgccttttgcgttctcaaaccttctctgtgc 1626
DB 1551 aagcgaagcgcatctctgtagtagtgccttttgcgttctcaaaccttctctgtgc 1610
OY 1627 atattacaagcattccccacagatacgttaactagcctcgttttgcatacagaag 1686
DB 1611 atattacaagcattccccacagatacgttaactagcctcgttttgcatacagaag 1670
OY 1687 gcaagctat 1694
DB 1671 gcaagctgt 1678

```

```

RESULT 11
AA083826
ID AA083826 standard; DNA; 5110 BP.
XX AC AA083826;
XX DT 04-OCT-1995 (first entry)
XX DE Plasmid pCAT10-1.
XX KM Plasmid pCAT10-1; bacteriophage T7; RNA polymerase;
XX mutant promoters: circular; cyclic; T7 promoter phi 10; ds.
XX OS Synthetic.
XX Key Location/Qualifiers
FH misc_feature 198..220
FT /tag= a
FT /note= "T7 promoter phi 10 prior to a promoterless
chloramphenicol acetyl transferase gene"
XX US5385834-A.
XX 31-JAN-1995.
XX 13-AUG-1993; 93US-0106433.
XX 13-AUG-1993; 93US-0106433.
XX (GBOR-) GEORGIA TECH RES CORP.
XX Ikeda RA:
XX WPI: 1995-081565/11.
XX New mutant form of T7 RNA polymerase - able to utilise a wide
PT range of T7 promoter sequences having mutations that inactivate
PT recognition by wild type enzyme
XX PS Disclosure: Columns 51-56; 79pp; English.
XX AA083826 is the plasmid pCAT10-1, it was used in the production of
CC the bacteriophage T7 RNA polymerase promoter phi 10.
XX CC Sequence 5110 BP; 1305 A; 1288 C; 1289 G; 1228 T; 0 other;
SQ
Query Match 22.5%; Score 404; DB 16; Length 5110;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY 1267 agagtcgacctgcagcagctgtggtccacactgacccatgcccgaactcagaag 1326
DB 1251 agaatttcctgcgagcagtagcgcggtgtccacactgacccatgcccgaactcagaag 1310
OY 1327 tgaacgcgctagcgcgagtagtgtgtggtgtcccccagtagaggaactgccc 1386
DB 1311 tgaacgcgctagcgcgagtagtgtgtggtgtcccccagtagaggaactgccc 1370
OY 1387 aggcataataaagcgaagctcagtcgaaagactggtgcttcttcttctgtgt 1446
DB 1371 aggcataataaagcgaagctcagtcgaaagactggtgcttcttcttctgtgt 1430
OY 1447 ttgtcgtggaacgctctctctgtagtagacaatcccgccggagcgagattggaactgtgcg 1506
DB 1431 ttgtcgtggaacgctctctctgtagtagacaatcccgccggagcgagattggaactgtgcg 1490
OY 1507 aagcaacgcccggaggtgtgcggaagcgcgccgcataaactgccaagcatcaatt 1566
DB 1491 aagcaacgcccggaggtgtgcggaagcgcgccgcataaactgccaagcatcaatt 1550

```

```
QY 1567 aagcagaaggcaltccctgacgaltgacctttttgctgttctacaacctcttcctgtcgtc 1626
      |||
DB 1551 aagcagaaggcaltccctgacgaltgacctttttgctgttctacaacctcttcctgtcgtc 1610
QY 1627 atatctacaagcattccccacagatacgttaactagctctgtttttgcatcagaaga 1686
      |||
DB 1611 atatctacaagcattccccacagatacgttaactagctctgtttttgcatcagaaga 1670
QY 1687 gcagctat 1694
      |||
DB 1671 gcagctgt 1678

RESULT 12
AAQ83827
ID AAQ83827 standard; DNA: 5110 BP.
XX
AC AAQ83827;
XX
DT 04-OCT-1995 (first entry)
XX
DE Plasmid PCM-T297.
XX
KM Plasmid PCM-T297; bacteriophage T7; RNA polymerase:
XX mutant -116 to T; T7 promoter phi 10; circular; cyclic; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Promoter 198..220
FT /tag= a
FT /note= "Inactive T7 promoter phi 10 mutant (-116 to
      T) prior to a promoterless chloramphenicol
      acetyl transferase gene"
XX
PN US5385834-A.
XX
PD 31-JAN-1995.
XX
PF 13-AUG-1993; 93US-0106433.
XX
PR 13-AUG-1993; 93US-0106433.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Ikeda RA:
XX
DR WPI: 1995-081565/11.
XX
PT New mutant form of T7 RNA polymerase - able to utilize a wide
PT range of T7 promoter sequences having mutations that inactivate
PT recognition by wild type enzyme
XX
PS Disclosure: Columns 57-62; 79pp; English.
XX
CC AAQ83827 is the plasmid PCM-T297, it was used in the production of
CC the bacteriophage T7 RNA polymerase mutant promoter phi 10 -116
CC to T substitution (claimed).
XX
SQ Sequence 5110 BP; 1305 A; 1288 C; 1288 G; 1229 T; 0 other;

Query Match 22.5%; Score 404; DB 16; Length 5110;
Best Local Similarity 96.5%; Pred. No. 4.9e-78;
Matches 413; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY 1387 aggcattcaataaagcgaagctcaatcgaagaactgggaccttgcgtttatctgtgtc 1446
      |||
DB 1371 aggcattcaataaagcgaagctcaatcgaagaactgggaccttgcgtttatctgtgtc 1430
QY 1447 ttgtcgttgaacgctctctctgtaagtaagaacaatccgcggaagcgtatggaactgtgcg 1506
      |||
DB 1431 ttgtcgttgaacgctctctctgtaagtaagaacaatccgcggaagcgtatggaactgtgcg 1490
QY 1507 aagcaaacgcccggagggttgccggcaggaacgcccgcataaactgccaagcattcaaat 1566
      |||
DB 1491 aagcaaacgcccggagggttgccggcaggaacgcccgcataaactgccaagcattcaaat 1550
QY 1567 aagcagaaggcaltccctgacgaltgacctttttgctgttctacaacctcttcctgtcgtc 1626
      |||
DB 1551 aagcagaaggcaltccctgacgaltgacctttttgctgttctacaacctcttcctgtcgtc 1610
QY 1627 atatctacaagcattccccacagatacgttaactagctctgtttttgcatcagaaga 1686
      |||
DB 1611 atatctacaagcattccccacagatacgttaactagctctgtttttgcatcagaaga 1670
QY 1687 gcagctat 1694
      |||
DB 1671 gcagctgt 1678

RESULT 13
AAQ83828
ID AAQ83828 standard; DNA: 5110 BP.
XX
AC AAQ83828;
XX
DT 04-OCT-1995 (first entry)
XX
DE Plasmid PCM-P1160.
XX
KM Plasmid PCM-P1160; bacteriophage T7; RNA polymerase:
XX mutant -9C to G; T7 promoter phi 10; circular; cyclic; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Promoter 198..220
FT /tag= a
FT /note= "Inactive T7 promoter phi 10 mutant (-9C to
      G) prior to a promoterless chloramphenicol
      acetyl transferase gene"
XX
PN US5385834-A.
XX
PD 31-JAN-1995.
XX
PF 13-AUG-1993; 93US-0106433.
XX
PR 13-AUG-1993; 93US-0106433.
XX
PA (GEOR-) GEORGIA TECH RES CORP.
XX
PI Ikeda RA:
XX
DR WPI: 1995-081565/11.
XX
PT New mutant form of T7 RNA polymerase - able to utilize a wide
PT range of T7 promoter sequences having mutations that inactivate
PT recognition by wild type enzyme
XX
PS Disclosure: Columns 63-70; 79pp; English.
XX
CC AAQ83828 is the plasmid PCM-P1160, it was used in the production of
CC the bacteriophage T7 RNA polymerase mutant promoter phi 10 -9C
CC to G substitution (claimed).
XX
SQ Sequence 5110 BP; 1300 A; 1291 C; 1293 G; 1226 T; 0 other;
```


